

Title: Proposed CR's from TSG SA WG4 to introduce SES to release 6 specifications

Source: TSG SA WG4

Agenda Item: 7.4.3

Document for: Information

Contact: Kari Jarvinen

Summary

The enclosed two CRs are proposed to introduce SES to release 6 specifications for packet switched conversational multimedia.

References

- [1] S4-040136 CR to TS 26.235 "Changes needed to add the DSR codec into the codec list"
- [2] S4-040131 CR to TS 26.236 "Changes needed to introduce SES payload formats"

3GPP TSG-S4#30 meeting
February 23 - 27, 2004, Malaga, Spain

Tdoc S4 (04)0131

CR-Form-v3

CHANGE REQUEST

⌘ **26.236 CR 010** ⌘ rev **1** ⌘ Current version: **5.4.0** ⌘

For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.

Proposed change affects: ⌘ (U)SIM ME/UE Radio Access Network Core Network

Title:	⌘ Introduction of the DSR Codec		
Source:	⌘ Motorola		
Work item code:	⌘ SRSES-Codec	Date:	⌘ 25.02.2004
Category:	⌘ B	Release:	⌘ REL-6
Use <u>one</u> of the following categories: F (essential correction) A (corresponds to a correction in an earlier release) B (Addition of feature), C (Functional modification of feature) D (Editorial modification) Detailed explanations of the above categories can be found in 3GPP TR 21.900.		Use <u>one</u> of the following releases: 2 (GSM Phase 2) R96 (Release 1996) R97 (Release 1997) R98 (Release 1998) R99 (Release 1999) REL-4 (Release 4) REL-5 (Release 5)	

Reason for change:	⌘ Additions for the DSR codec
Summary of change:	⌘ Changes needed to introduce SES payload formats
Consequences if not approved:	⌘ The DSR feature is not introduced

Clauses affected:	⌘ 2, 3.2, 4, 5.4	
Other specs Affected:	<input type="checkbox"/> Other core specifications <input type="checkbox"/> Test specifications <input type="checkbox"/> O&M Specifications	⌘
Other comments:	⌘	

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The present document specifies the codec specific RTP protocol details applying to packet switched conversational multimedia applications within the 3GPP IM Subsystem.

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
 - 1 presented to TSG for information;
 - 2 presented to TSG for approval;
 - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

Introduction

The present document contains a specification for required protocol usage within 3GPP specified Conversational Packet Switched Multimedia Services [5] which is based IP Multimedia Subsystem (IM Subsystem). IM Subsystem as a subsystem includes specifically the conversational IP multimedia services, whose service architecture, call control and media capability control procedures have been defined in 3GPP TS 24.229 [7], and are based on the 3GPP adopted version of IETF Session Initiated Protocol (SIP) [1].

In conversational packet switched multimedia service depends on IM Subsystem. The individual media types are independently encoded and packetized to appropriate separate Real Time Protocol (RTP) packets. These packets are then transported end-to-end inside UDP datagrams over real-time IP connections that have been negotiated and opened between the terminals [or between a terminal and a server](#) during the SIP call as specified in 3GPP TS 24.229 [7].

The UEs operating within IM Subsystem need to provide encoding/decoding of the derived codecs, and perform corresponding packetization/depacketization functions. Logical bound between the media streams is handled in the SIP session layer, and inter-media synchronization in the receiver is handled with the use of RTP time stamps.

1 Scope

The present document introduces the required protocols for packet switched conversational multimedia applications within 3GPP IP Multimedia Subsystem. Visual and sound communications are specifically addressed. The intended applications are assumed to require low-delay, real-time functionality.

The present document describes the required protocol related elements for 3G PS multimedia terminal:

- required SDP signalling regarding the media type bit rate, packet size, packet transport frequency;
- usage of RTP payload for media types;
- bandwidth adaptation;
- QoS negotiation.

The present document is applicable, but not limited, to packet switched video telephony.

The applicability of the present document to GERAN is FFS.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

- [1] IETF RFC 2543: "SIP: Session Initiation Protocol".
- [2] IETF RFC 2327: "SDP: Session Description Protocol".
- [3] IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications", Schulzrinne H. et al, July 2003.
- [4] IETF RFC 3551: "RTP Profile for Audio and Video Conferences with Minimal Control", Schulzrinne H. and Casner S., July 2003.
- [5] 3GPP TS 26.235: "Packet switched conversational multimedia applications; Default codecs".
- [6] 3GPP TS 24.228: "Signalling flows for the IP multimedia call control based on SIP and SDP; stage 3".
- [7] 3GPP TS 24.229: "IP multimedia call control protocol based on SIP and SDP".
- [8] 3GPP TS 23.228: "IP Multimedia Ssbsystem (IMS); Stage 2".
- [9] 3GPP TS 23.107: "Quality of Service (QoS) concept and architecture".
- [10] 3GPP TS 23.207: "End to end quality of service concept and architecture".
- [11] 3GPP TS 23.060: "General Packet Radio Service (GPRS); Service description; Stage 2".
- [12] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; General description".

- [13] 3GPP TS 26.090: "AMR speech Codec; Transcoding Functions".
- [14] 3GPP TS 26.073: "AMR speech Codec; C-source code".
- [15] 3GPP TS 26.104: "ANSI-C code for the floating-point Adaptive Multi-Rate AMR speech codec".
- [16] 3GPP TS 26.171 (Release 5): "AMR speech codec, wideband; General description".
- [17] 3GPP TS 26.190 (Release 5): "Mandatory Speech Codec speech processing functions AMR Wideband speech codec; Transcoding functions".
- [18] 3GPP TS 26.201 (Release 5): "AMR speech codec, wideband; Frame structure".
- [19] [IETF RFC 3267: "RTP payload format and file storage format for the Adaptive Multi-Rate \(AMR\) Adaptive Multi-Rate Wideband \(AMR-WB\) audio codecs", March 2002.](#)
- ~~3GPP TS 26.235: "Packet switched conversational multimedia applications; Default codecs". Annex B: "RTP payload format and storage format for AMR and AMR-WB audio".~~
- [20] ITU-T Recommendation H.263: "Video coding for low bit rate communication".
- [21] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)".
- [22] ISO/IEC 14496-2 (1999): "Information technology - Coding of audio-visual objects - Part 2: Visual".
- [23] IETF RFC 3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams".
- [24] ITU-T Recommendation H.263 (annex X): "Annex X: Profiles and levels definition".
- [25] 3GPP TS 26.235: "Packet Switched Conversational Multimedia Applications; Default Codecs ". Annex C: "ITU-T H.263 MIME media type registration".
- [26] ITU-T Recommendation T.140 (1998): "Protocol for multimedia application text conversation" (with amendment 2000).
- [27] IETF RFC 2793: "RTP Payload for Text Conversation".
- [28] IETF RFC 3556: "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) bandwidth", Casner S., July 2003.
- [29] [RTP Payload Formats for European Telecommunications Standards Institute \(ETSI\) European Standard ES 202 050, ES 202 211, and ES 202 212 Distributed Speech Recognition Encoding draft-ietf-avt-rtp-dsr-codecs-00.txt](#)

CR Editor's note: The above document cannot be formally referenced until it is published as an RFC.

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following term and definition applies:

3G PS multimedia terminal: terminal based on IETF SIP/SDP internet standards modified by 3GPP for purposes of 3GPP IM Subsystem services

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AMR	Adaptive MultiRate codec
DSR	Distributed Speech Recognition
IETF	Internet Engineering Task Force
IM Subsystem	Internet protocol Multimedia Subsystem
ITU-T	International Telecommunications Union-Telecommunications
RFC	IETF Request For Comments
RTPCP	RTP Control Protocol
RTP	Real-time Transport Protocol
SDP	Session Description Protocol
SES	Speech Enabled Service
SIP	Session Initiation Protocol

4 General

3G PS multimedia terminals provide real-time video, audio, [SES](#) or data, in any combination, including none, over 3GPP IM Subsystem. Terminals are based on IETF defined multimedia protocols SIP, SDP, RTP and RTCP. Communication may be either 1-way or 2-way. Such terminals may be part of a portable device or integrated into an automobile or other non-fixed location device. They may also be fixed, stand-alone devices; for example, a video telephone or kiosk. Multimedia terminals may also be integrated into PCs and workstations.

[In the case of SES then uplink communication is from the terminal to a server containing speech recognition.](#)

In addition, interoperation with other types of multimedia telephone terminals, such as 3G-324M may be possible, however in such case a media gateway functionality supporting 3G-324M - IM Subsystem interworking will be required within or outside the IM subsystem.

Figure 1 presents the user plane protocol stack of a 3G PS conversational multimedia terminal explaining the transport of different media types and QoS reports.

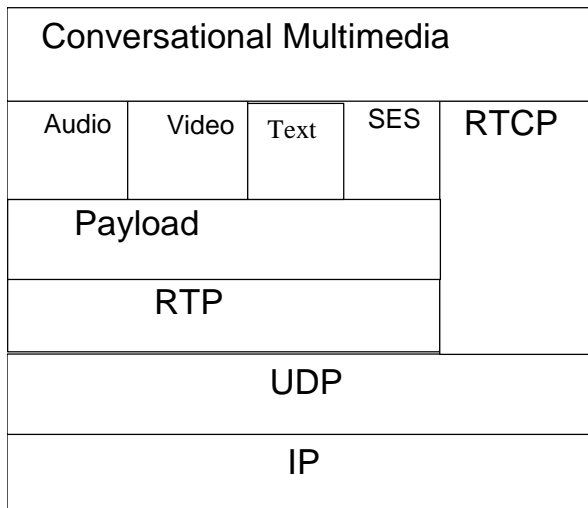


Figure 1 – User plane protocol stack for 3G PS conversational multimedia terminal

5 Media type requirements

Media type RTP payload usage is specified in this clause. The media types and corresponding codecs are specified in 3GPP TS 26.235 [5]. The continuous media type RTP payloads are mapped to RTP packets according to IETF RTP Profile for Audio and Video Conferences with Minimal Control in RFC 3551 [4].

5.1 Audio

5.1.1 RTP session description parameters

The IETF AMR and AMR-WB RTP payload format [19] offers different options. Here is the list of options and how they should be used by the transmitter. The receiver shall at least support the options as they are listed:

- the bandwidth efficient operation shall be used,
- only one speech frame shall be encapsulated in each RTP packet,
- the multi-channel session shall not be used,
- interleaving shall not be used,
- internal CRC shall not be used.

5.2 Video

Video packets should not be large to allow better error resilience and to minimize the transmission delay in conversational service. The size of each packet shall be kept smaller than 512 bytes.

5.3 Real time text

Real time text media type RTP payload format for ITU-T Recommendation T.140 is specified in [27]. Redundant transmission provided by the RTP payload format is recommended in error prone channel.

5.4 SES

[The RTP payload for the DSR codec and AMR or AMR-WB used for SES are specified in \[29, 19\].](#)

6 Call control

Functional requirements for call control are specified in 3GPP TS 23.228 [8].

The required signalling functions are specified in 3GPP TS 24.228 [6] and call control protocols in 3GPP TS 24.229 [7].

QoS authorization issues and interworking with the IM subsystem in general are covered in 3GPP TS 23.207 [10].

7 Bearer control

The media control is based on declaration of terminal media capability sets in SDP part of appropriate SIP messages. The usage of bearer bandwidth can be effectively controlled by adjusting the media type encoder bit rates.

7.1 Bandwidth

The bandwidth information of each media type shall be carried in SDP messages in both session and media type level during codec negotiation, session establishment and resource reallocation. Note that for RTP based applications, 'b=AS:' gives the RTP "session bandwidth" (including UDP/IP overhead) as defined in section 6.2 of [3].

The bandwidth for RTCP traffic shall be described using the "RS" and "RR" SDP bandwidth modifiers at media level, as specified by [28]. Therefore, a conversational multimedia terminal shall include the "b=RS:" and "b=RR:" fields in

SDP, and shall be able to interpret them. There shall be a limit on the allowed RTCP bandwidth for a session signalled by the terminal. This limit is defined as follows:

- 4000 bps for the RS field (at media level);
- 3000 bps for the RR field (at media level).

7.2 QoS negotiation

The QoS architecture and concept is specified in 3GPP TS 23.107 [9]. The end-to-end QoS framework involving GPRS and UMTS is specified in 3GPP TS 23.207 [10]. The applicable general QoS mechanism and service description for the GPRS in GSM and UMTS is specified in 3GPP TS 23.060 [11].

7.3 RTP receiver

The RTP receiver implementation and functionality including lost and delayed packet processing as well as jitter buffer is out of scope of the present document.

Annex A (informative): Optional enhancements

This annex is intended for informational purposes only. This is not an integral part of the present document.

A.1 Video enhancements

This clause gives informative recommendations for the video media type control.

The SDP attributes regarding the video frame rate and the quality of media encoding should be used to ensure good video service. The recommended usage of these attributes are FFS.

`a=framerate:<frame rate>` describes the maximum video frame rate attribute in frames/second. Fractional values of `<frame rate>` are allowed.

`a=quality:<quality>` describes the quality of media encoding attribute, where the `<quality>` is a value in [0..10] with 10 indicating the best quality.

Annex B (informative): Mapping of SDP parameters to UMTS QoS parameters

This clause gives recommendations for mapping of SDP parameters in UMTS QoS parameters for conversational multimedia applications. Different use cases will be considered. Each use case generates an example QoS profile parameters table (with values for IPv4 and IPv6 addressing). The values indicated are derived by applications' QoS requirements, and may not be fulfilled by the network. In the parameters for guaranteed and maximum bit rates a granularity of 1 kbps is assumed for bearers up to 64 kbps, as defined in the TS 24.008. Therefore the "Ceiling" function is used for up-rounding fractional values, wherever needed. In addition, the same specification defines a granularity of 10 bytes for the Maximum SDU sizes values. This is taken into account in the computation of this field in the QoS profile.

Use case 1 – Voice over IP

This use case includes the scenario in which two conversational multimedia terminals establish a bi-directional Voice over IP (VoIP) connection for speech communication, using the AMR or AMR-WB codecs with the same bit rate in both uplink and downlink directions.

For example an AMR VoIP stream encoded at 12.2 kbps, with one speech frame encapsulated into an RTP packet, would yield IP packets of the following size (using the mandated bandwidth efficient mode):

20 (IPv4) + 8 (UDP) + 12 (RTP) + 32 (AMR RTP payload) = 72 bytes, or

40 (IPv6 with no extension headers) + 8 (UDP) + 12 (RTP) + 32 (AMR RTP payload) = 92 bytes.

The gross bit rate including uncompressed RTP/UDP/IPv4 headers would be 28.8 kbps. The value in the b=AS media level parameter would be 29. The gross bit rate including uncompressed RTP/UDP/IPv6 headers would be 36.8 kbps. The value in the b=AS media level parameter would be 37.

To determine the Maximum SDU size parameter we should consider the maximum packet size that can be generated with a speech codec. This is exactly that generated by a AMR-WB stream at 23.85 kbps packetized in bandwidth efficient mode and with 1 speech frame per packet. Considering uncompressed RTP/UDP/IPv6 headers, the maximum packet size is 121 bytes.

The QoS profile would be set then using the following parameters:

Table B.1: QoS profile for AMR VoIP at 12.2 kbps

QoS parameter	Parameter value	Comment
Delivery of erroneous SDUs	No	
Delivery order	No	To minimize delay in the access stratum. The application should take care of eventual packet reordering
Traffic class	Conversational	
Maximum SDU size	130 bytes	10 bytes granularity. The RTP packet size might change the maximum SDU size limitation [tbc]
Guaranteed bitrate for downlink	SDP media bw in DL + 2.5% * (SDP media bw in DL+ SDP media bw in UL) = Ceil(30.45)=31 kbps (for the IPv4 case) Ceil(38.85)=39 kbps (for the IPv6 case)	
Maximum bit rate for downlink	Ceil(30.45)=31 kbps (for the IPv4 case) Ceil(38.85)=39 kbps (for the IPv6 case)	
Guaranteed bitrate for uplink	SDP media bw in UL + 2.5% * (SDP media bw in UL+ SDP media bw in DL) = Ceil(30.45)=31 kbps (for the IPv4 case) Ceil(38.85)=39 kbps (for the IPv6 case)	
Maximum bit rate for uplink	Ceil(30.45)=31 kbps (for the IPv4 case) Ceil(38.85)=39 kbps (for the IPv6 case)	
Residual BER	10^{-5}	16 bit CRC
SDU error ratio	$7 \cdot 10^{-3}$	
Traffic handling priority	Not used in Conversational traffic class	
Transfer delay	100 ms	
SDU format information	Not used	
Allocation/retention priority	Subscribed allocation/retention priority	Not relevant for the application
Source statistics descriptor	"Speech"	

In some cases, multiple AMR or AMR-WB rates are available, and rate control techniques allow to switch between different modes based on the received speech quality. For example, if the available AMR mode set is {4.75, 10.2, 12.2} kbps, the set of gross bit rates are:

AMR 4.75 kbps: 21.6 kbps (including RTP/UDP/IPv4 headers). [SDP b=AS parameter would be 22].

AMR 10.2 kbps: 26.8 kbps (including RTP/UDP/IPv4 headers). [SDP b=AS parameter would be 27].

AMR 12.2 kbps: 28.8 kbps (including RTP/UDP/IPv4 headers). [SDP b=AS parameter would be 29].

In case of IPv6 addressing, the gross bit rates are:

AMR 4.75 kbps: 29.6 kbps (including RTP/UDP/IPv6 headers). [SDP b=AS parameter would be 30].

AMR 10.2 kbps: 34.8 kbps (including RTP/UDP/IPv6 headers). [SDP b=AS parameter would be 35].

AMR 12.2 kbps: 36.8 kbps (including RTP/UDP/IPv6 headers). [SDP b=AS parameter would be 37].

The maximum bit rate is set to the highest mode of the codec. However, the procedure on how to choose the guaranteed bit rate when several codec rates are available is to be defined. Here we provide an example QoS profile in which the guaranteed speech quality is at least that of 10.2 kbps AMR for both uplink and downlink directions, while the non-guaranteed maximum quality is that of 12.2 kbps for both uplink and downlink directions.

Table B.2: QoS profile for AMR VoIP at 3 bit rates with rate control

QoS parameter	Parameter value	Comment
Delivery of erroneous SDUs	No	
Delivery order	No	To minimize delay in the access stratum. The application should take care of eventual packet reordering
Traffic class	Conversational	
Maximum SDU size	130 bytes	10 bytes granularity. The RTCP packet size might change the maximum SDU size limitation [tbc]
Guaranteed bitrate for downlink	SDP media bw in DL + 2.5% * (SDP media bw in DL+ SDP media bw in UL) = Ceil(28.35)=29 kbps (for the IPv4 case) Ceil(36.75)=37 kbps (for the IPv6 case)	Guaranteed quality 10.2 kbps
Maximum bit rate for downlink	SDP media bw in DL + 2.5% * (SDP media bw in DL+ SDP media bw in UL) = Ceil(30.35)=31 kbps (for the IPv4 case) Ceil(38.85)=39 kbps (for the IPv6 case)	Non-guaranteed quality 12.2 kbps
Guaranteed bitrate for uplink	SDP media bw in UL+ 2.5% * (SDP media bw in UL+ SDP media bw in DL) = Ceil(28.35)=29 kbps (for the IPv4 case) Ceil(36.75)=37 kbps (for the IPv6 case)	Guaranteed quality 10.2 kbps
Maximum bit rate for uplink	SDP media bw in UL + 2.5% * (SDP media bw in UL+ SDP media bw in DL) = Ceil(30.35)=31 kbps (for the IPv4 case) Ceil(38.85)=39 kbps (for the IPv6 case)	Non-guaranteed quality 12.2 kbps
Residual BER	10^{-5}	16 bit CRC
SDU error ratio	$7 \cdot 10^{-3}$	
Traffic handling priority	Not used in Conversational traffic class	
Transfer delay	100 ms	
SDU format information	Not used	
Allocation/retention priority	Subscribed allocation/retention priority	Not relevant for the application
Source statistics descriptor	"Speech"	

Use case 2 – Unidirectional video

This use case includes the scenario in which two conversational multimedia terminals establish a uni-directional video connection, using the H.263 or MPEG-4 codecs.

The video codec in this example has a bitrate of 36 kbps, with RTP payload packets of 75 bytes (excluding payload header which is, for example, 2 bytes). The sending terminal would produce IP packets of the following size:

20 (IPv4) + 8 (UDP) + 12 (RTP) + 77 (video RTP payload+payload header) = 117 bytes, or

40 (IPv6 with no extension headers) + 8 (UDP) + 12 (RTP) + 77 (video RTP payload+payload header) = 137 bytes.

The gross bit rate including uncompressed RTP/UDP/IPv4 headers would be 56.2 kbps. The value in the b=AS media level parameter would be 57. The gross bit rate including uncompressed RTP/UDP/IPv6 headers would be 65.8 kbps. The value in the b=AS media level parameter would be 66.

The maximum video packet size is limited to 512 bytes in section 5.2. This value is fine if transmission occurs over the UMTS Iu interface. However, in order to avoid SNDCP fragmentation of packets over the GERAN Gb interface (where the default size for LLC data field (=SNDCP frame) is 500 bytes) the maximum IP packet size is 500 – 4 (unacknowledged mode SNDCP header) = 496 bytes. Therefore, the maximum size of a video packet is 496 – 60 (RTP/UDP/IPv6 uncompressed headers) = 436 bytes (including RTP payload header). 400 bytes is a safer value.

The QoS profile of the receiving terminal would be set then using the following parameters:

Table B.3: QoS profile for unidirectional video at 36 kbps

QoS parameter	Parameter value	Comment
Delivery of erroneous SDUs	No	
Delivery order	No	To minimize delay in the access stratum. The application should take care of eventual packet reordering
Traffic class	Conversational	
Maximum SDU size	500 bytes	10 bytes granularity
Guaranteed bitrate for downlink	SDP media bw in DL + 2.5% * (SDP media bw in DL) = Ceil(58.43)=59 kbps (for the IPv4 case) Ceil(67.65)=68 kbps (for the IPv6 case)	
Maximum bit rate for downlink	Equal or higher than guaranteed bit rate	
Guaranteed bitrate for uplink	2.5% * (SDP media bw in DL) = Ceil(1.43)=2 kbps (for the IPv4 case) Ceil(1.65)=2 kbps (for the IPv6 case)	For RTCP
Maximum bit rate for uplink	Equal or higher than guaranteed bit rate	
Residual BER	10 ⁻⁵	16 bit CRC
SDU error ratio	10 ⁻³	
Traffic handling priority	Not used in Conversational traffic class	
Transfer delay	250 ms	
SDU format information	Not used	
Allocation/retention priority	Subscribed allocation/retention priority	Not relevant for the application
Source statistics descriptor	"Unknown"	

Use case 3 – Video telephony

This use case includes the scenario in which two conversational multimedia terminals establish a bi-directional speech/video connection, using the AMR/AMR-WB and H.263/MPEG-4 codecs at the same bit rates in uplink and downlink directions.

The video codec in this case has a bitrate of 28 kbps, with RTP payload packets of 250 bytes (excluding payload header which is, for example, 2 bytes). The total video bit rate is 32.7 kbps (including RTP/UDP/IPv4 headers). The value in the b=AS media level parameter would be 33. For IPv6 addressing, the total video bit rate is 34.9 kbps (including RTP/UDP/IPv6 headers). The value in the b=AS media level parameter would be 35.

In the same bearer there is an AMR stream at 10.2 kbps with 1 frame encapsulated per RTP packet using the bandwidth efficient mode. The total voice bit rate is 26.8 kbps (including RTP/UDP/IPv4 headers). The value in the b=AS media level parameter would be 27. For IPv6 addressing, the total voice bit rate is 34.8 kbps (including RTP/UDP/IPv6 headers). The value in the b=AS media level parameter would be 35.

The total media bit rate is $28+10.2=38.2$ kbps. The total session bit rate is $33+27=60$ kbps for IPv4 addressing, and $35+35=70$ kbps for IPv6 addressing.

The terminal would produce IP packets of the following size:

AMR: 20 (IPv4) + 8 (UDP) + 12 (RTP) + 27 (AMR RTP payload) = 67 bytes (or 87 bytes for IPv6 with no extension headers).

Video: 20 (IPv4) + 8 (UDP) + 12 (RTP) + 252 (video RTP payload+payload header) = 292 bytes (or 312 bytes for IPv6 with no extension headers).

The same considerations done in Use Case 2 about the maximum packet sizes apply also for this use case.

The QoS profile of the videotelephony terminal would be set then using the following parameters:

Table B.4: QoS profile for videotelephony at 38.2 kbps

QoS parameter	Parameter value	Comment
Delivery of erroneous SDUs	No	
Delivery order	No	To minimize delay in the access stratum. The application should take care of eventual packet reordering
Traffic class	Conversational	
Maximum SDU size	500 bytes	10 bytes granularity
Guaranteed bitrate for downlink	SDP media bw in DL for AMR + $2.5\% * (\text{SDP media bw in DL for AMR} + \text{SDP media bw in UL for AMR}) +$ SDP media bw in DL for video + $2.5\% * (\text{SDP media bw in DL for video} + \text{SDP media bw in UL for video})$ = Ceil(63.0)=63 kbps (for the IPv4 case) = Ceil(73.3)=74 kbps (for the IPv6 case)	
Maximum bit rate for downlink	Equal or higher than guaranteed bit rate	
Guaranteed bitrate for uplink	SDP media bw in UL for AMR + $2.5\% * (\text{SDP media bw in UL for AMR} + \text{SDP media bw in DL for AMR}) +$ SDP media bw in UL for video + $2.5\% * (\text{SDP media bw in UL for video} + \text{SDP media bw in DL for video})$ = Ceil(63.0)=63 kbps (for the IPv4 case) = Ceil(73.3)=74 kbps (for the IPv6 case)	
Maximum bit rate for uplink	Equal or higher than guaranteed bit rate	
Residual BER	10^{-5}	16 bit CRC
SDU error ratio	10^{-3}	
Traffic handling priority	Not used in Conversational traffic class	
Transfer delay	100 ms	
SDU format information	Not used	
Allocation/retention priority	Subscribed allocation/retention priority	Not relevant for the application
Source statistics descriptor	"Unknown"	

In case of usage of separate PDP contexts for the speech and video streams, the speech stream QoS profile parameters are set similarly to use case 1, while the video stream QoS profile parameters are set similarly to use case 2 (but considering that the video flow is bi-directional and considering possibly the same UMTS bearer transfer delay constraints for both media).

Annex C (informative): Change history

Change history							
Date	TSG #	TSG Doc.	CR	Rev	Subject/Comment	Old	New
2002-03	15	SP-020074			Version 2.0.0 presented for approval	2.0.0	5.0.0
2002-12	18	SP-020695	001	2	QoS profile parameters for conversational multimedia applications	5.0.0	5.1.0
2002-12	18	SP-020695	002	1	Clarification on SDP session bandwidth parameter	5.0.0	5.1.0
2003-03	19	SP-030092	003	2	SDP bandwidth modifier for RTCP bandwidth	5.1.0	5.2.0
2003-03	19	SP-030092	004		Correction on QoS profile parameters for conversational multimedia applications	5.1.0	5.2.0
2003-06	20	SP-030219	005		Examples of QoS profiles for conversational multimedia applications	5.2.0	5.3.0
2003-09	21	SP-030449	006		Correction of obsolete RTP references	5.3.0	5.4.0
2003-09	21	SP-030449	007	1	Correction of wrong reference	5.3.0	5.4.0

3GPP TSG-S4#30 meeting
February 23 - 27, 2004, Malaga, Spain

Tdoc S4 (04)0136

CR-Form-v3

CHANGE REQUEST

⌘ 26.235 CR 006 ⌘ rev 2 ⌘ Current version: 6.00 ⌘

For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.

Proposed change affects: ⌘ (U)SIM ME/UE Radio Access Network Core Network

Title: ⌘ Introduction of the DSR Codec

Source: ⌘ Motorola

Work item code: ⌘ SRSES-Codec **Date:** ⌘ 27.01.2004

Category: ⌘ B **Release:** ⌘ REL-6

Use one of the following categories:

- F (essential correction)
- A (corresponds to a correction in an earlier release)
- B (Addition of feature),
- C (Functional modification of feature)
- D (Editorial modification)

Detailed explanations of the above categories can be found in 3GPP TR 21.900.

Use one of the following releases:

- 2 (GSM Phase 2)
- R96 (Release 1996)
- R97 (Release 1997)
- R98 (Release 1998)
- R99 (Release 1999)
- REL-4 (Release 4)
- REL-5 (Release 5)

Reason for change: ⌘ Additions for the DSR codec

Summary of change: ⌘ Changes needed to add the DSR codec into the codec list

Consequences if not approved: ⌘ The DSR feature is not introduced

Clauses affected: ⌘ 2, 3.2, 6.5, 9.1 and 9.2

Other specs Affected: ⌘ Other core specifications ⌘ Test specifications O&M Specifications

Other comments: ⌘

1 Scope

The present document introduces the set of default codecs for packet switched conversational multimedia applications within 3GPP IP Multimedia Subsystem. Visual and sound communication are specifically addressed. The intended applications are assumed to require low-delay, real-time functionality.

The present document is applicable, but not limited, to PS video telephony.

The applicability of this specification to GERAN is FFS.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

- [1] IETF RFC 3261: "SIP: Session Initiation Protocol".
- [2] IETF RFC 2327: "SDP: Session Description Protocol".
- [3] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)".
- [4] IETF RFC 1889: "RTP: A Transport Protocol for Real-Time Applications".
- [5] IETF RFC 3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams".
- [6] ITU-T Recommendation H.263: "Video coding for low bit rate communication".
- [7] 3GPP TS 26.110: "Codec for Circuit Switched Multimedia Telephony Service; General Description".
- [8] 3GPP TS 26.111: "Codec for Circuit Switched Multimedia Telephony Service; Modifications to H.324".
- [9] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; General description".
- [10] 3GPP TS 26.090: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Transcoding functions".
- [11] 3GPP TS 26.073: "Adaptive Multi-Rate (AMR); ANSI C source code".
- [12] 3GPP TS 26.104: "ANSI-C code for the floating-point AMR speech codec".
- [13] ISO/IEC 14496-2 (1999): "Information technology - Coding of audio-visual objects - Part 2: Visual".
- [14] 3GPP TS 24.228: "Signalling flows for the IP multimedia call control based on SIP and SDP".
- [15] 3GPP TS 24.229: "IP Multimedia Call Control Protocol based on SIP and SDP".
- [16] 3GPP TS 26.171 (Release 5): "AMR speech codec, wideband; General description".

- [17] 3GPP TS 26.190 (Release 5): "Mandatory Speech Codec speech processing functions AMR Wideband speech codec; Transcoding functions".
- [18] 3GPP TS 26.201 (Release 5): "AMR speech codec, wideband; Frame structure".
- [19] ITU-T Recommendation H.263 (annex X): "Annex X, Profiles and levels definition".
- [20] 3GPP TS 23.228: "IP multimedia subsystem; stage 2".
- [21] 3GPP TS 23.107: "QoS Concept and Architecture".
- [22] 3GPP TS 23.207: "End to end quality of service concept and architecture".
- [23] 3GPP TS 23.060: "General Packet Radio Service (GPRS); Service description; Stage 2".
- [24] IETF RFC 2793: "RTP Payload for Text Conversation".
- [25] ITU-T Recommendation T.140 (1998): "Protocol for multimedia application text conversation" (with amendment 2000).
- [26] 3GPP TS 26.101: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Frame Structure".
- [27] IETF RFC 2119: "Key words for use in RFCs to Indicate Requirement Levels".
- [28] 3GPP TS 26.093: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Source Controlled Rate operation".
- [29] 3GPP TS 46.060: "Enhanced Full Rate (EFR) speech transcoding".
- [30] TIA/EIA -136-Rev.A, part 410 - "TDMA Cellular/PCS – Radio Interface, Enhanced Full Rate Voice Codec (ACELP). Formerly IS-641. TIA published standard, 1998".
- [31] ARIB, RCR STD-27H, "Personal Digital Cellular Telecommunication System RCR Standard".
- [32] IETF draft-westberg-realtime-cellular-01.txt, "Realtime Traffic over Cellular Access Networks".
- [33] IETF draft-larzon-udplite-03.txt, "The UDP Lite Protocol".
- [34] 3GPP TS 26.092: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Comfort noise aspects".
- [35] IETF RFC 3267: "RTP payload format and file storage format for the Adaptive Multi-Rate (AMR) Adaptive Multi-Rate Wideband (AMR-WB) audio codecs", March 2002.
- [36] IETF RFC 2833: "RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals", May 2000.
- [37] [3GPP TS 26.243: "Draft TS Software documentation for fixed-point DSR Extended Front End"](#)
- [38] [RTP Payload Formats for European Telecommunications Standards Institute \(ETSI\) European Standard ES 202 050, ES 202 211, and ES 202 212 Distributed Speech Recognition Encoding draft-ietf-avt-rtp-dsr-codecs-00.txt](#)

CR Editor's note: The above document cannot be formally referenced until it is published as an RFC.

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply.:

3G PS multimedia terminal: terminal based on IETF SIP/SDP internet standards modified by 3GPP for purposes of 3GPP packet switched network based multimedia telephony

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AMR	Adaptive MultiRate codec
DSR	Distributed Speech Recognition
IETF	Internet Engineering Task Force
IM Subsystem	Internet protocol Multimedia Subsystem
ITU-T	International Telecommunications Union-Telecommunications
RFC	IETF Request For Comments
RTCP	RTP Control Protocol
RTP	Real-time Transport Protocol
SDP	Session Description Protocol
SES	Speech Enabled Services
SIP	Session Initiated Protocol

4 General

3G PS multimedia terminals provide real-time video, audio, or data, in any combination, including none, over 3GPP IM Subsystem. Terminals are based on IETF defined multimedia protocols SIP, SDP, RTP and RTCP. Communication may be either 1-way or 2-way. Such terminals may be part of a portable device or integrated into an automobile or other non-fixed location device. They may also be fixed, stand-alone devices; for example, a video telephone or kiosk. Multimedia terminals may also be integrated into PCs and workstations.

In addition, interoperation with other types of multimedia telephone terminals, such as 3G-324M may be possible, however in such case a media gateway functionality supporting 3G-324M - IM Subsystem interworking will be required within or outside the IM subsystem.

5 System overview

The present document describes the required codec related elements for 3G PS multimedia terminal:

- mandatory and optional codecs for 3G PS multimedia terminal;
- media encapsulation and decapsulation rules for each mandatory and optional codec.

6 Functional requirements

SIP protocol itself does not mandate any codecs. Standardisation of mandatory codecs does not prevent the use of other codecs that can be signalled using the SDP protocol. 3G PS multimedia terminals shall be able to use the same audio and video codecs applied in 3G-324M [8]. This will ensure the interoperability with 3G circuit switched multimedia telephony.

6.1 Audio

3G PS multimedia terminals offering audio communication shall support AMR narrowband speech codec [9], [10], [\[11+\]](#) to [12]. This is the mandatory speech codec.

The AMR wideband speech codec shall be supported when the 3G PS multimedia terminal supports wideband speech working at 16 kHz sampling frequency [16].

The usage of telephone-event media format is recommended for DTMF.

6.2 Video

3G PS multimedia terminals offering video communication shall support ITU-T recommendation H.263 [6] baseline. This is the mandatory video codec.

H.263 [19] version 2 Interactive and Streaming Wireless Profile (Profile 3) Level 10 should be supported. This is an optional video codec.

ISO/IEC 14496-2 [13] (MPEG-4 Visual) Simple Profile at Level 0 should be supported. This is an optional video codec.

6.3 Real time text

3G PS multimedia terminals offering real time text conversation should support ITU-T Recommendation T.140 [25] Text Conversation presentation coding.

6.4 Interactive and background data

SIP signalling offers initialisation of packet switched interactive or background class reliable data services as well. However specification of such data services are outside the scope of the present document.

6.5 Speech Enabled Service

3G PS multimedia terminals offering speech enabled services should support the DSR Extended Advanced Front-end codec [37]

Speech enabled services may also be supported with AMR or AMR-WB audio codecs, however it is noted that there is substantial performance advantage from DSR [ref to S4-040145].

7 Call control

Functional requirements for call control are specified in 3GPP TS 23.228 [20].

The required signalling functions are specified in 3GPP TS 24.228 [14] and call control protocols in 3GPP TS 24.229 [15].

8 Bearer control

The media control is based on declaration of terminal media capability sets in SDP part of appropriate SIP messages.

Relation of application level SDP signalling and radio access bearer assignment is defined outside the present document. The QoS architecture and concept for WCDMA and GERAN is specified in 3GPP TS 23.107 [21]. The end-to-end QoS framework involving GPRS and UMTS is specified in 3GPP TS 23.207 [22]. The applicable general QoS mechanism and service description for the GPRS in GSM and UMTS is specified in 3GPP TS 23.060 [23].

9 Multimedia stream encapsulation

9.1 MIME media types

The terminal shall declare the mandatory and any optional media streams using the codec specific MIME media types in the associated SDP syntax. The MIME media types for the mandatory and optional codecs shall be according to the corresponding types registered by IANA.

- AMR narrowband speech codec MIME media type as specified in annex B.
- AMR wideband speech codec MIME media type is specified in annex B.

- H.263 [6] video codec MIME media type is specified in annex C.
- MPEG-4 visual simple profile level 0 MIME media type as specified in RFC 3016 [5].
- ITU-T Recommendation T.140 [25] Text Conversation MIME media type as specified by RFC 2793 [24].
- Telephone-event MIME media type as specified by RFC 2833 [36].
- [DSR MIME media type as specified in draft-ietf-avt-rtp-dsr-codecs-00.txt \[38\]](#)

9.2 RTP payload

RTP payload formats specified by IETF shall be used for real time media streams.

RTP payload format for the AMR narrowband speech codec is specified in annex B.

RTP payload format for the AMR wideband speech codec is specified in annex B.

RTP payload format for the ITU-T Recommendation H.263 [6] video codec is specified in IETF RFC 2429 [3].

RTP payload format for the MPEG-4 visual simple profile level 0 is specified in IETF RFC 3016 [5].

RTP payload format for the ITU-T Recommendation T.140 [25] text conversation coding is specified in IETF RFC 2793 [24].

RTP payload format for the telephone-event is specified in IETF RFC 2833 [36].

[RTP payload format for the DSR Extended Advanced Front-end is specified in draft-ietf-avt-rtp-dsr-codecs-00.txt \[38\]](#)

Annex A (informative): Information on optional enhancements

This annex is intended for informational purposes only. This is not an integral part of the present document.

A.1 Video

This clause gives recommendations for the video codec implementations within 3G PS multimedia terminals.

Regardless of which specific video codec standard is used, all video decoder implementations should include basic error concealment techniques. These techniques may include replacing erroneous parts of the decoded video frame with interpolated picture material from previous decoded frames or from spatially different locations of the erroneous frame. The decoder should aim to prevent the display of substantially corrupted parts of the picture. In any case, it is recommended that the terminal should tolerate *every* possible bitstream without catastrophic behaviour (such as the need for a user-initiated reset of the terminal).

3G PS terminal video encoders and decoders are recommended to support the 1:1 pixel format (square format).

A.1.1 H.263 video codec

H.263 was approved as a standard in 1996. Since then, version 2 and version 3 enhancing version 1 have been approved in 1998 and 2000 respectively. As of today, H.263 contains an extensive set of mandatory and optional coding tools. H.263 [6] annex X (going to be approved in 2001) defines codec profiles for various target environments.

The Baseline Profile (Profile 0) stands for H.263 with no optional modes of operation. It includes the basic coding tool set common in modern video coding standards. It provides simple means to insert resynchronisation points within the video bitstream, and, therefore, it enables recovery from erroneous or lost data.

The Version 2 Interactive and Streaming Wireless Profile (Profile 3) provides enhanced compression efficiency when compared to the Baseline Profile. Moreover, it provides enhanced error resilience for delivery to wireless devices. Specifically, Profile 3 includes the following optional coding modes:

- 1) Advanced INTRA Coding (annex I). Use of this mode improves the compression efficiency for INTRA macroblocks (whether within INTRA pictures or predictively-coded pictures);
- 2) Deblocking Filter (annex J). A deblocking filter improves image quality by reducing blocking artifacts. When compared to deblocking filtering performed as a postprocessing operation, the Deblocking Filter Mode reduces the amount of required memory, as no additional picture memory is needed for the filtered images. This mode also includes the four-motion-vector-per-macroblock feature and picture boundary extrapolation for motion compensation, both of which can further improve compression efficiency;
- 3) Slice Structured Mode (annex K). This mode provides a flexible mechanism to insert resynchronisation points within the video bitstream for recovery from erroneous or lost data.
- 4) Modified Quantisation (annex T). This mode enables flexible quantiser control that can be used in sophisticated bit-rate control algorithms. In addition, it improves chrominance fidelity.

[FFS]

A.2 Audio

[FFS]

A.3 Text

Use of the redundancy coding variant specified in RFC 2793 [24] is recommended for error resilience.

Annex B (normative): AMR and AMR-WB RTP payload and MIME type registration

The AMR and AMR-WB speech codec RTP payload, storage format and MIME type registration are specified in [35].

Annex C (normative): ITU-T H.263 MIME media type registration

NOTE: The intention is to replace this normative annex with the IETF RFC defining the H.263 [6] video codec MIME media type registration when the RFC is available.

H.263 video codec MIME media type is specified as follows:

MIME media type name: video;

MIME subtype name: H263-2000;

Required parameters: None;

Optional parameters:

- profile: H.263 profile number, in the range 0 through 8, specifying the supported H.263 annexes/subparts;
- level: Level of bitstream operation, in the range 0 through 99, specifying the level of computational complexity of the decoding process. When profile and level parameters are not specified, Baseline Profile (Profile 0) Level 10 are the default values.

The profile and level specifications can be found in [19]. Note that the RTP payload format for H263-2000 is the same as for H263-1998 published in RFC 2429 [3], but additional annexes/subparts are specified along with the profiles and levels.

Annex D (informative): Change history

Change history							
Date	TSG SA#	TSG Doc.	CR	Rev	Subject/Comment	Old	New
03-2001	11	SP-010095			Version for Release 4		4.0.0
06-2001	12	SP-010309	001		Update of AMR-NB and AMR-WB RTP payload	4.0.0	4.1.0
	12				Release 4 version withdrawn		
06-2001	12	SP-010378	002		Applicability of TS 26.235 to GERAN FFS	4.0.0	5.0.0
	12				TS 26.235 approved at TSG-SA#12 for Release 5		5.0.0
03-2002	15	SP-020089	003	2	Update of AMR & AMR-WB RTP payload format	5.0.0	5.1.0
03-2002	15	SP-020154	004		Corrections of references to obsolete SIP RFC 2543 IETF specification	5.0.0	5.1.0
06-2003	20	SP-030218	005	1	Handling of DTMF in IMS	5.1.0	6.0.0